Project KLAM – Outline

1. Introduction: describe project scenario – what did we hope to accomplish/learn?

# Data Set Description

# Explanation of dataset: domain, source

# Size

# What Data Structure for Data Storage

# Data Preprocessing

# Data cleaning

# Outlier removal

# Other data sanitization/wrangling

# Data Analysis/Data Processing (at least 4)

# Analysis/Processing #1

# Analysis/Processing #2

# Analysis/Processing #3

# Analysis/Processing #4

# Analysis/Processing #5+ (if necessary/relevant(

# Testing: Describe Test-driven Development / Unit Tests

# Test-driven development

# Unit tests

# Explanation of Results & Conclusions

# How results can be used by others: relevance/significance

# Ways to improve/expand/add/remove functionality

# Presentation Skills & Video, Live & Written

# Video presentation (how to split?):

# Celeste

# John

# Sarai

# Jamie

# Live session presentation (limit 4-7 minutes)

# Written presentation (see Explanation of Results step above)

# Peer Evaluations – to be assigned

# Extra Credit Topics

# Web scraping

# User-interaction to retrieve/display/analyze results

# User-interaction to obtain additional data

# Advanced queries/manipulate the data in another way